AFRL-HE-AZ-TP-2002-0011

UNITED STATES AIR FORCE RESEARCH LABORATORY

ASSESSMENT OF THE VIRTUAL ENVIRONMENT SAFE-FOR-MAINTENANCE TRAINER (VEST)

BRENDA M. WENZEL

ANNA R. CASTILLO

Air Force Research Laboratory
Warfighter Training Research Division
6030 South Kent Street
Mesa, AZ 85212-6061

GERRY BAKER

363 TRS/TTW
520 Missile Road
Sheppard Air Force Base, TX 76311

SEPTEMBER 2002

AIR FORCE MATERIEL COMMAND
AIR FORCE RESEARCH LABORATORY
HUMAN EFFECTIVENESS DIRECTORATE
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (FROM - TO)
01-10-2002	Final	xx-04-2002 to xx-08-2002

4. TITLE AND SUBTITLE
Assessment of the Virtual Environment Safe-for-Mainenance Trainer (VEST)
Unclassified

5a. CONTRACT NUMBER
F416245-97-D-5000

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
Wenzel, Brenda M. ; Author
Castillo, Anna R. ; Author
Baker, Gerry ; Author

5d. PROJECT NUMBER
1123

5e. TASK NUMBER
B0

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Air Force Research Laboratory
Warfighter Training Research Division
6030 South Kent Street
Mesa, AZ85212-6061

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL-HE-AZ-TP-2002-0011

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
Air Force Research Laboratory
Warfighter Training Research Division
6030 South Kent Street
Mesa, AZ85212-6061

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL; AFRL/HEA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES
Air Force Research Laboratory Technical Monitor: Dr Brenda M. Wenzel, AFRL/HEAS, 480-988-6561, x-225, DSN 474-6225

14. ABSTRACT

The 363d Training Squadron (TRS), Sheppard AFB TX, is assessing the training effectiveness and usability of their Virtual Environment Safe-for-maintenance Trainer (VEST). VEST is an immersive virtual reality (VR) environment that provides demonstrations, drills, and checks on performance for F-15E model (a) cockpit switch familiarization, (b) weapons station identification, and (c) external aircraft safe-for-maintenance procedures. The assessment includes performance data and apprentices? reactions to the training experience. Results from the initial assessment of VEST provide considerable evidence for the value in training aircraft maintenance technician apprentices with VR technology. Although training in a VR environment may be a new experience for apprentices, their performance scores reveal acceptable levels of learning. Results from the continuous assessment support the earlier finding that VEST provides effective training and experience that would otherwise not be afforded F-15E aircraft maintenance apprentice technicians. Apprentice comments reveal an overall positive learning experience with VEST. Their comments also reveal a negative perceptual experience. Improvements need to be made to VEST to minimize negative aspects of training in virtual environments. The negative comments indicated ?fixable? problems, i.e., instructional aspects, graphics quality, joystick difficulties, monotonous synthetic voice, and general usability.

15. SUBJECT TERMS
Cockpit switchology; Immersive virtual reality; Open-ended survey; Safe-for-maintenance; Training; Training assessment; Training effectiveness; VEST; Virtual Environment Safe-for-maintenance Trainer; Virtual reality

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	Public Release	26	Liz, Casey liz.casey@williams.af.mil
Unclassified	Unclassified	Unclassified			19b. TELEPHONE NUMBER International Area Code Area Code Telephone Number 480988-6561 DSN 474-6188

NOTICES

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of STINFO exchange.

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

BRENDA M. WENZEL
Project Scientist

DEE H. ANDREWS
Technical Advisor

CURTIS J. PAPKE, Colonel, USAF
Chief, Warfighter Training Research Division

Copies of this report may be requested from:

Defense Technical Information Center
8725 John J. Kingman Road, Suite 0944
Ft. Belvoir, VA 22060-6218
http://stinet.dtic.mil

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
September 2002	Final	Apr 02 to Aug 02

4. TITLE AND SUBTITLE

Assessment of the Virtual Environment Safe-for-Maintenance Trainer (VEST)

5a. CONTRACT NUMBER
F41624-97-D-5000 T0 17

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)

Brenda M. Wenzel
Anna R. Castillo
*Gerry Baker

5d. PROJECT NUMBER
1123

5e. TASK NUMBER
B0

5f. WORK UNIT NUMBER
01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

*363d TRS/TTW
520 Missile Road
Sheppard Air Force Base, TX 76311

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-HE-AZ-TP-2002-0011

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Air Force Research Laboratory Technical Monitor: Dr Brenda M. Wenzel, AFRL/HEAS, 480-988-6561, x-225, DSN 474-6225

14. ABSTRACT

The 363d Training Squadron (TRS), Sheppard AFB TX, is assessing the training effectiveness and usability of their Virtual Environment Safe-for-maintenance Trainer (VEST). VEST is an immersive virtual reality (VR) environment that provides demonstrations, drills, and checks on performance for F-15E model (a) cockpit switch familiarization, (b) weapons station identification, and (c) external aircraft safe-for-maintenance procedures. The assessment includes performance data and apprentices' reactions to the training experience. Results from the initial assessment of VEST provide considerable evidence for the value in training aircraft maintenance technician apprentices with VR technology. Although training in a VR environment may be a new experience for apprentices, their performance scores reveal acceptable levels of learning. Results from the continuous assessment support the earlier finding that VEST provides effective training and experience that would otherwise not be afforded F-15E aircraft maintenance apprentice technicians. Apprentice comments reveal an overall positive learning experience with VEST. Their comments also reveal a negative perceptual experience. Improvements need to be made to VEST to minimize negative aspects of training in virtual environments. The negative comments indicated "fixable" problems, i.e., instructional aspects, graphics quality, joystick difficulties, monotonous synthetic voice, and general usability.

15. SUBJECT TERMS

Cockpit switchology; Immersive virtual reality; Open-ended survey; Safe-for-maintenance; Training; Training assessment; Training effectiveness; VEST; Virtual Environment Safe-for-maintenance Trainer; Virtual reality

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Ms Liz Casey
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED	26	**19b. TELEPHONE NUMBER** *(include area code)* 480.988.6561 x- 188 DSN 474-6188

ASSESSMENT OF THE VIRTUAL ENVIRONMENT SAFE-FOR-MAINTENANCE TRAINER (VEST)

EXECUTIVE SUMMARY

The 363d Training Squadron (TRS), at Sheppard AFB TX, does not have a full-scale, F-15E model aircraft in the hangar for training aircraft maintenance apprentices. All apprentices receive baseline training on the F-15C model aircraft. At some point in training, apprentices are given duty assignments to bases with specific F-15 aircraft models. Those assigned to maintain the F-15E are required to complete 21 lessons in the Virtual Environment Safe-for-maintenance Trainer (VEST). VEST is an immersive virtual reality (VR) environment that provides demonstrations, drills, and checks on performance for F-15E model (a) cockpit switch familiarization, (b) weapons station identification, and (c) external aircraft safe-for-maintenance procedures.

The 363d TRS conducted an initial assessment of VEST and is continually assessing VEST. The assessments include performance data and apprentices' reactions to the training experience. Results from both assessments are reported here. Results from the initial assessment of VEST provide considerable evidence for the value in training aircraft maintenance technician apprentices with VR technology. Although training in a VR environment may be a new experience for apprentices, their performance scores reveal acceptable levels of learning. The following statements describe the results. Overall, VEST is well received by apprentices as a training experience. Some apprentices feel VEST had a positive impact on their classroom training. Apprentices recommend expanding VR to other types of training. Training with VR increases apprentice confidence interacting with an F-15E model. Lesson length is not satisfactory for many apprentices. Various adverse physical effects are experienced in the VR environment. VEST needs improvements to its instructional features, physical configuration, peripherals (joystick, headset) and synthetic voice.

Valuable information on the effectiveness and acceptance of VEST is revealed through the continuous assessment data. The data support the earlier finding that VEST provides effective training and experience that would otherwise not be afforded F-15E model aircraft maintenance apprentice technicians. Apprentice comments and performance scores reveal an overall positive learning experience with VEST. Their comments also reveal a negative perceptual experience. No relationships were found between performance and comment characteristics. This finding suggests any negative reaction to VEST did not influence overall performance.

Improvements need to be made to VEST to minimize negative aspects of training in virtual environments. Learning in virtual environments does not have to be painful. The negative comments indicate "fixable" problems—instructional aspects, graphics quality, joystick difficulties, monotonous synthetic voice, and general usability.

The assessment data allow researchers to begin considering the next-generation specifications for an enhanced VR system by basing improvements to the new system on apprentice performance and reactions and considering leading-edge VR technologies. A collaborative effort is underway between the Air Force Research Laboratory and the Air Education and Training Command to reengineer VEST to a Generalized Operations Simulation Environment (GOSE) for aircraft maintenance. GOSE is conceptualized as a scalable, modular, immersive VR training system comprised of common PC-based hardware and software.

ASSESSMENT OF THE VIRTUAL ENVIRONMENT SAFE-FOR-MAINTENANCE TRAINER (VEST)

INTRODUCTION

The 363d Training Squadron (TRS), Sheppard AFB TX, does not have a full-scale F-15E model aircraft in the hangar for training aircraft maintenance apprentices. All apprentices receive baseline training on the F-15C model aircraft. At some point in training, apprentices are given duty assignments to bases with specific F-15 models. Those assigned to maintain the F-15E are required to complete 21 lessons in the Virtual Environment Safe-for-maintenance Trainer (VEST). A prerequisite of VEST is successful completion of objectives in baseline training on the F-15C.

The value of VEST is its low-cost replacement of the actual aircraft, thus, providing apprentices with training opportunities they would not otherwise have prior to deployment. VEST provides apprentices with contextualized, 3-dimensional, interactive experiences with the E-model front and rear crew stations, weapons stations, and ground safe-for-maintenance.

There are two VEST stations in operation at the 363d TRS (see Figure 1). It takes approximately two and one half hours to complete VEST. The trainer provides demonstrations, drills, and checks on performance. Apprentices are given two attempts to achieve a 70% performance score to pass a drill. VEST lessons cover the following topics:

- Introduction to virtual reality (VR),
- Forward crew station panels and switchology,
- Aft crew station panels and switchology,
- E-model weapons stations,
- Particular E-model safety devices, and
- Aircraft ground safe-for-maintenance.

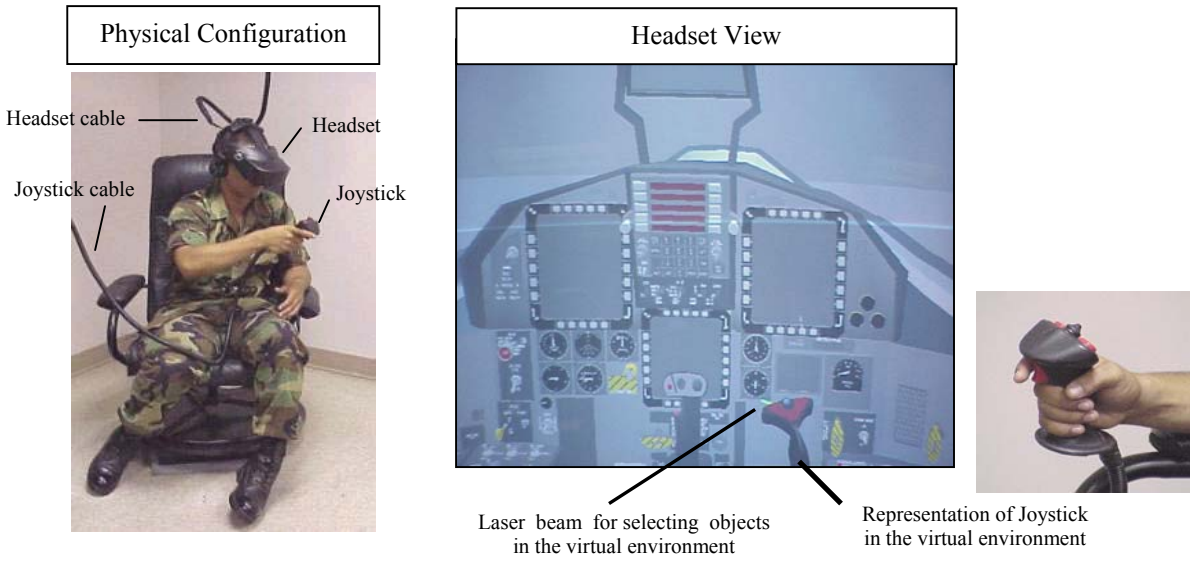

Figure 1. VEST System in Operation

The 363d TRS conducted an initial assessment of VEST and are continually assessing VEST. The assessments include apprentice performance in the VR environment and their reactions to the training experience. Results to date are reported here.

INITIAL ASSESSMENT

Data Collection

Data from 64 aircraft maintenance apprentices who participated in the initial assessment of VEST are described here. Performance data was automatically collected by VEST across 11 drills. The apprentices completed an open-ended written survey after training with VEST. The survey consisted of 13 items ranging from previous experience with VR to strong and weak points of their experience with VEST. Performance data from 12 apprentices was not captured by the system so their data have been excluded from the analyses reported here.

Performance Data

Table 1 contains average performance scores across the VEST drills broken out by number of tries. Safe-for-maintenance drill scores were missing for three more apprentices; their data were only excluded from the overall performance score. The overall performance score was calculated as the average score across the 11 drills. The average overall performance score for the remaining 49 apprentices was 88.4%. Only one apprentice (66.6%) had an overall performance score below 70%. Practice had a noticeable impact on the *aircraft safe for maintenance drill* where those who required a second try at attaining a passing score significantly outscored those who passed on a single try ($\underline{M}_{1st\ try}$ = 70.6, $\underline{M}_{2nd\ try}$ = 86.9, $\underline{t}_{(47)}$ = 6.1; $\underline{p}$value = .0001).

Table 1. Average Performance Scores on VEST Drills by Number of Tries (n = number of apprentices)

	Number of Tries			Number of Tries	
Drill	**1**	**2**	**Drill**	**1**	**2**
Left console forward cockpit switch drill	84.3 (n=42)	72.6 (n=10)	Left console aft cockpit switch drill	93.0 (n=46)	80.0 (n=6)
Main console forward cockpit switch drill	85.0 (n=44)	78.5 (n=8)	Main console aft cockpit switch drill	86.4 (n=47)	86.0 (n=5)
Right console forward cockpit switch drill	90.2 (n=49)	70.3 (n=3)	Right console aft cockpit switch drill	85.0 (n=46)	80.0 (n=6)
Review aft cockpit drill	100 (n=17)	82.9 (n=35)			
Weapons station drill	91.0 (n=49)	83.7 (n=3)	Safety device drill	94.5 (n=49)	89.5 (n=2)
Aircraft safe for maintenance intro drill	100 (n=50)	---	Aircraft safe for maintenance drill	70.6 (n=16)	86.9 (n=33)

Table 2 reports the number of apprentices who did not pass a drill even after two tries. Thirty-five percent of apprentices scored below 70% on at least one drill. Of those, one apprentice missed a passing score on four of the switchology drills and three others missed passing scores on two drills--a switchology and safe-for-maintenance.

Table 2. Number of Apprentices Scoring Below 70 Percent Across Drills

Drill	No.	Drill	No.
Left console forward cockpit switch drill	4	Left console aft cockpit switch drill	0
Main console forward cockpit switch drill	2	Main console aft cockpit switch drill	1
Right console forward cockpit switch drill	1	Right console aft cockpit switch drill	1
Review aft cockpit drill	5		
Weapons station drill	0	Safety device drill	0
Aircraft safe for maintenance intro drill	0	Aircraft safe for maintenance drill	9

Open-ended Survey

Table 3 presents 12 of the 13 items from an open-ended survey that apprentices completed following training. These 12 items asked for explanations/comments from apprentices based on their initial responses to the items. The missing survey item in Table 3, "Was enough time provided for you to complete the virtual reality lesson and the normal classroom training?" did not require further explanation and therefore is not listed in the table. Two raters made categorical judgments of respondents' explanations/comments. Cohen's index "kappa" was used to derive the degree of agreement in category assignment. Average interrater reliability (.90) was acceptably high across the 12 open-ended items. Individual "kappas" are found in Table 4 below. Prior to calculating response percentages within categories, raters came to agreement on category assignments. Comment categories and response percentages are presented next.

Table 3. Open-ended Survey Items, Number of Apprentice Comments Collected and Interrater Reliabilities for Coded Comments

Survey Item	Number of Comments	Cohen's "kappa"
Have you ever experienced any type of virtual reality games or training prior to this time? (If so, provide a brief explanation of your experience)	17	.81
Were you able to complete the lessons without any adverse physical effects? (Explain)	47	1.00
Did it provide you with enough reality for you to enter an F15-E aircraft cockpit and feel safe in doing so? (Explain)	39	.70
Was the computer voice easy to understand?	36	.88
Did you have any difficulty with the headset viewer, the joystick or other components of the equipment?	39	.78
Were the lessons too long or short? (Explain)	49	.92
What effect, if any, did the virtual reality training have on the normal classroom training?	64	.83
Would you recommend virtual reality training for other types of training? (If so, in what areas?)	47	.96
If you could, what would you change in the virtual world?	63	.94
STRONG POINTS	59	.89
WEAK POINTS	66	1.00
ADDITIONAL COMMENTS	31	1.00

Prior experience with virtual reality games or VR training. Seventy-one percent of apprentices reported no experience with VR prior to completing VEST and 29% reported prior VR experience. The range of prior VR experience is shown in Figure 2.

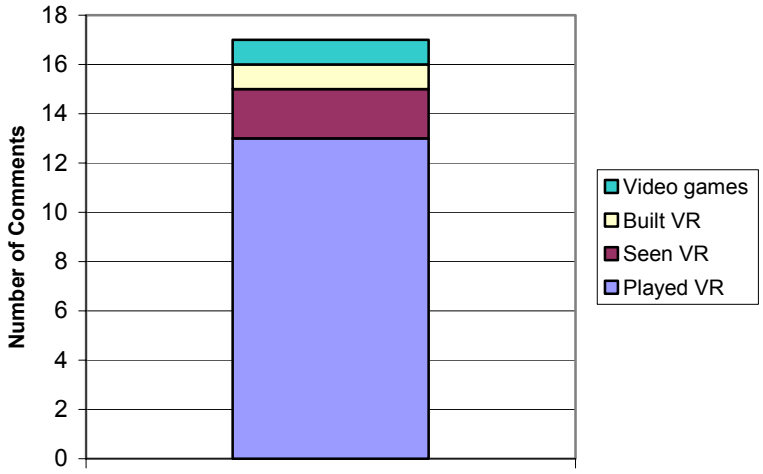

Figure 2. Prior Experience with Virtual Reality

Lessons were completed without any adverse physical effects. Twenty-nine percent of apprentices reported no adverse physical effects associated with completing VEST. Fifty-eight percent of apprentices reported "YES" they did have adverse physical effects to VEST. Thirteen percent of apprentices commented on their experience without stating "NO" or "YES" to the item. Figure 3 presents a breakout of comments within categories.

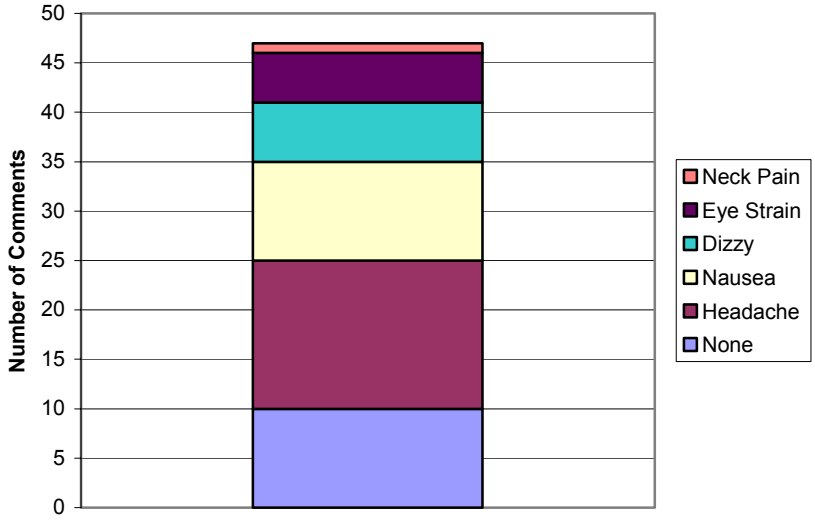

Figure 3. Physical Reactions to the VEST Experience

Lessons provided sufficient reality to feel safe entering the F-15E cockpit. Eighty-six percent of apprentices reported "YES" the lessons provided sufficient reality. Twelve percent of apprentices reported "NO" and 2% of apprentices did not respond. Some responses are questionable as to whether apprentices fully understood the question, for example, "*The cockpit and nomenclatures are familiar to me. I'm better off for having done this course*."; "*The hands-on feel made it familiar"; and "I think I could safe the F-15E without hesitation*." With that in mind, Figure 4 presents a breakout of comments within categories.

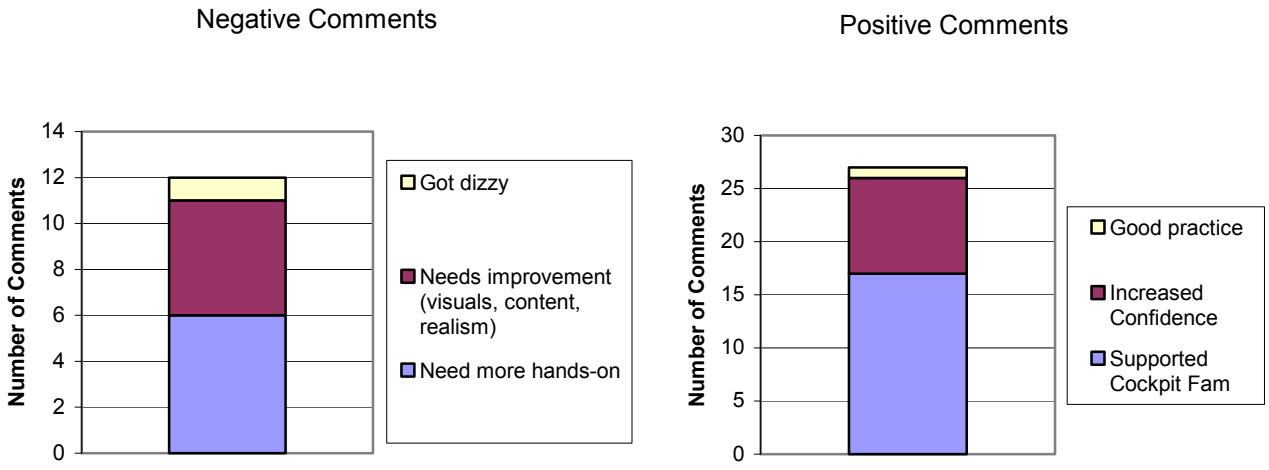

Figure 4. Comments Regarding Sufficient Reality to Feel Safe in the Cockpit

Computer voice was easy to understand. Fifteen percent of apprentices reported "NO" the synthetic voice used in VEST was not easily understood. Eight-five percent of apprentices reported a *qualified* "YES" the voice was easy to understand, with 11% of the "YES" respondents only commenting on the item. Figure 5 presents a breakout of comments within categories.

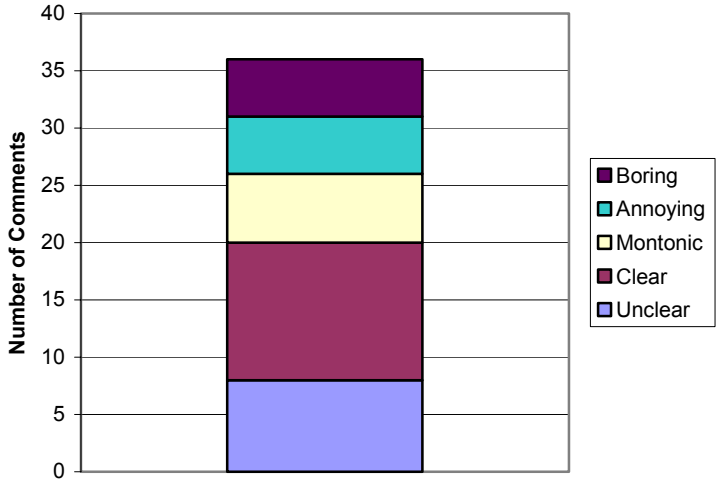

Figure 5. Reactions to the Synthetic Voice

Experienced difficulty with the headset viewer, the joystick, or other components of the equipment. Thirty-three percent of apprentices reported a *qualified* no difficulties with peripherals. Thirty-eight percent of apprentices reported "YES" they experienced difficulties with peripherals. Twenty-nine percent of apprentices only commented on the item. Figure 6 presents a breakout of comments within categories.

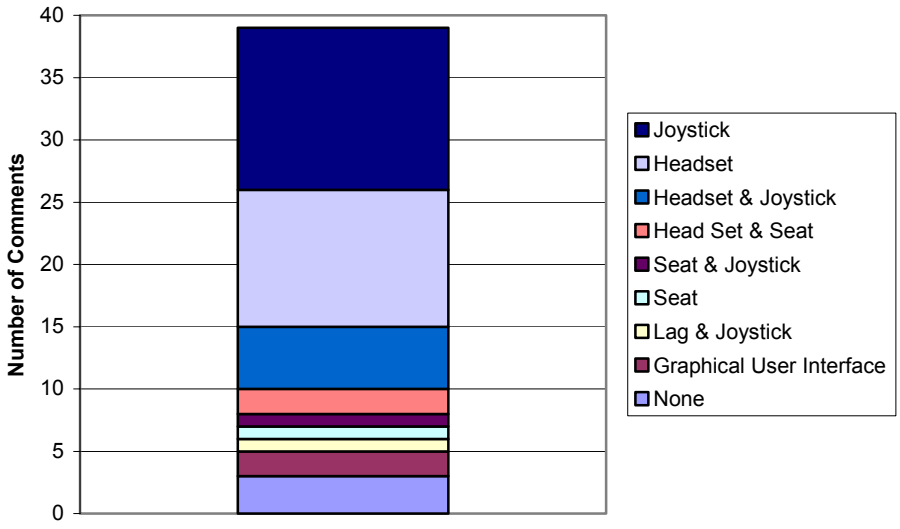

Figure 6. Difficult Peripherals

Lessons were too long or too short. Sixty-five percent of apprentices reported the length of the VEST lessons to be "OK." Seventeen percent of apprentices reported them to be too long, 8% reported lessons were too short, and 4% reported some lessons were long and some were short. Six percent of apprentices did not respond to the item. Figure 7 presents a breakout of comments within categories.

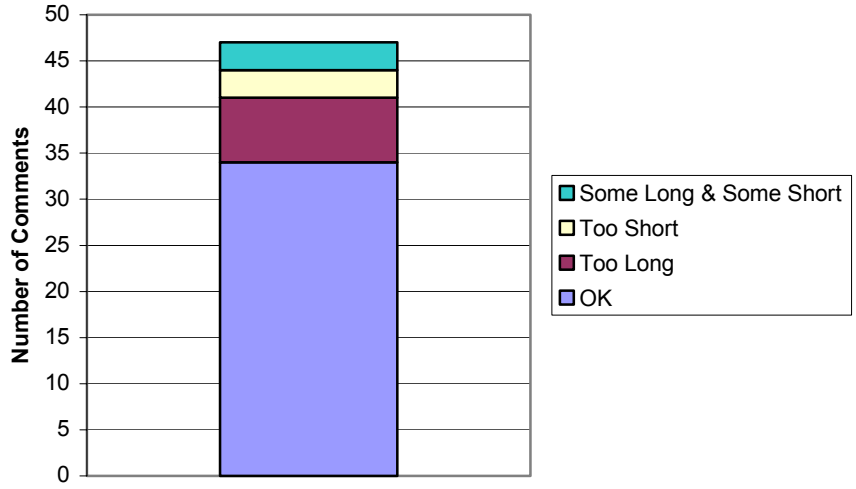

Figure 7. Perceived Length of Lessons

Enough time was provided to complete the virtual reality lessons and normal classroom training. Ninety-two percent of apprentices reported "YES" there was enough time to complete both VEST and classroom training. The remaining 8% only commented on the item: "I had plenty of time to complete the virtual reality lesson. That's why I understand the objective pretty well."; "It took up time that could have been spent otherwise. Many could have been shorter as a whole?????"; "The lessons were too long but not long enough for the information to sink in."; "Some lessons told you to click on the objects, but would continue on without you."

Effect virtual reality training had on classroom training. Forty-six percent of apprentices reported a positive effect of VEST on classroom training; while 54% of respondents reported VEST had no effect on classroom training. Figure 8 presents a breakout of comments within categories.

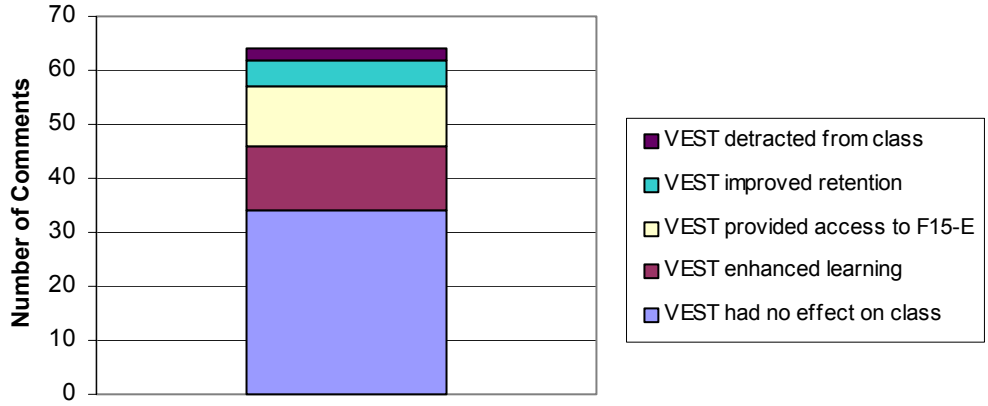

Figure 8. Effect of Virtual Reality Training on Classroom Training

Recommendations for use of virtual reality training in other types of training. Twenty-three percent of apprentices either made no recommendations for other training applications of VR or stated, "Don't know." Four percent of apprentices did not respond to the item. Of the remaining apprentices, 52% reported "YES" they would recommend VR for other types of training and 21% only commented. Figure 9 presents a breakout of comments within categories.

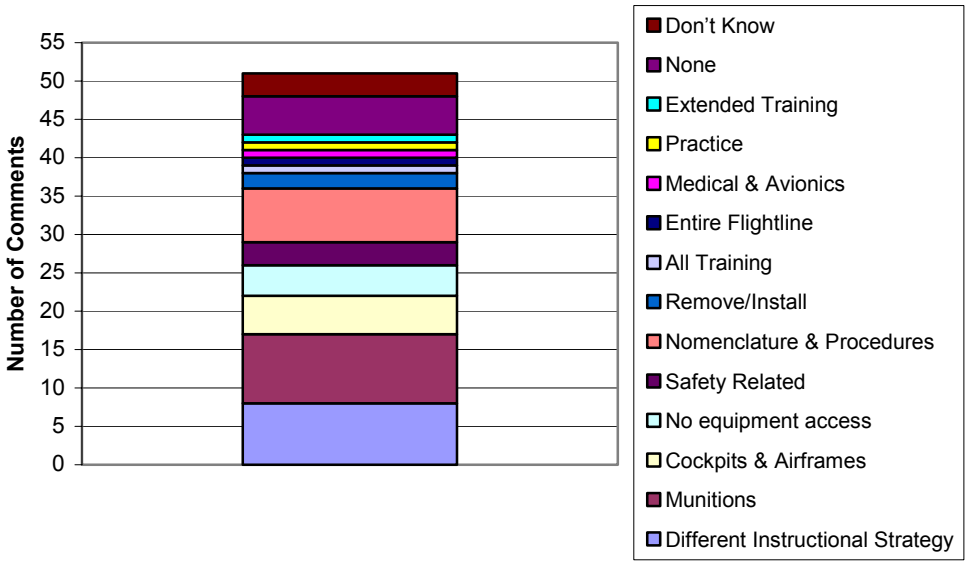

Figure 9. Recommendations for Use of Virtual Reality in Other Types of Training

Suggested changes to the virtual world. Seventy-nine percent of apprentices suggested changes to VEST. Of those suggestions, 8% addressed instructional design issues, 12% addressed graphical user interface (GUI) issues, 17% addressed the synthetic voice, 19% addressed realism, and 23% addressed hardware issues. Nineteen percent of apprentices reported nothing needed to be changed and 2% did not respond to the item. Figure 10 presents a breakout of comments across the *GUI*, *instruction*, *realism*, and *peripheral* categories.

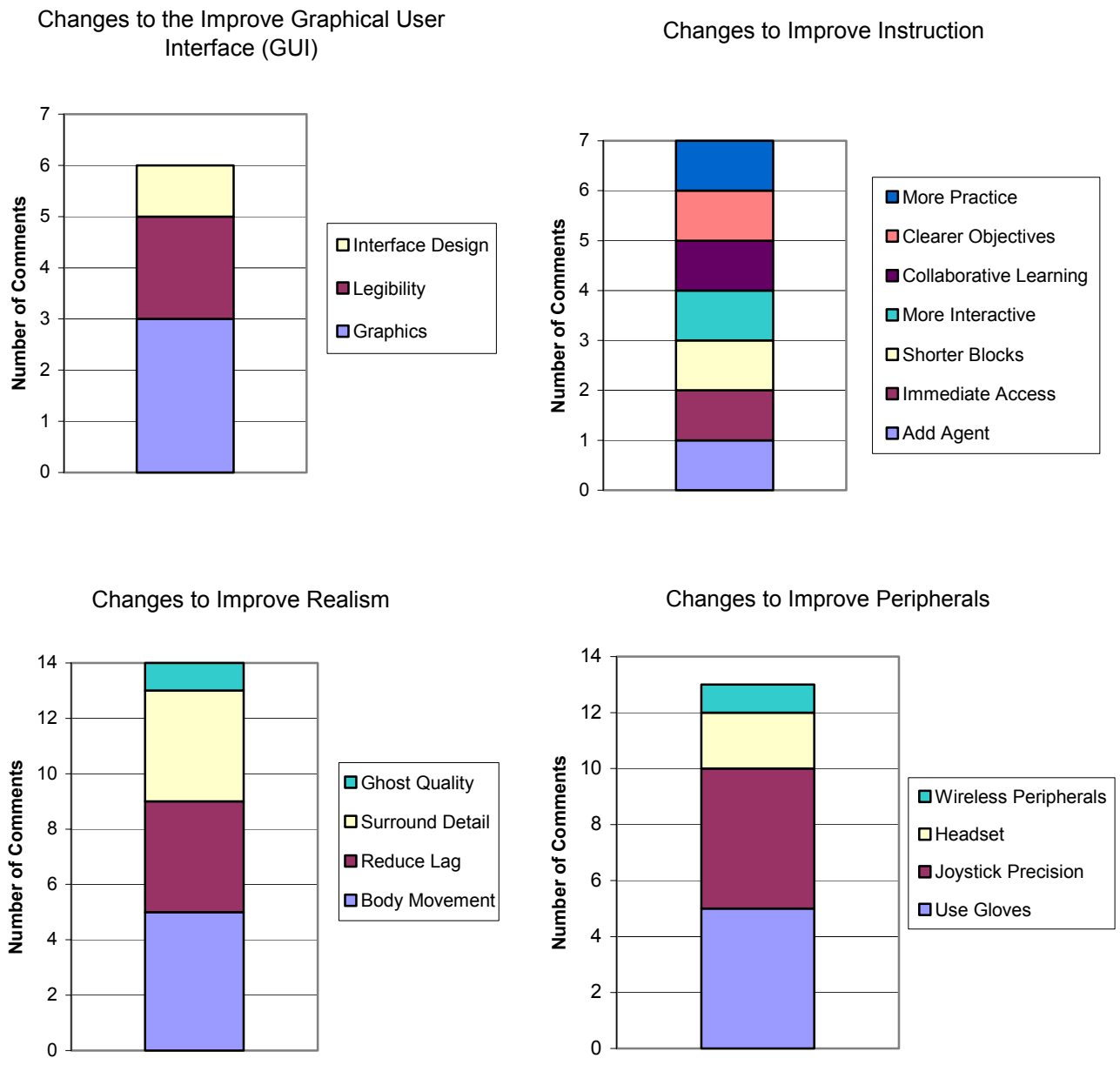

Figure 10. Suggested Changes to Improve Virtual Environments

Strong points of VEST. Eighty-five percent of apprentices commented on strong points of VEST. Of those, 6% concerned the graphical representations, 11% concerned the instructional value, 15% concerned the VR trainer itself, 15% concerned realism of the VR system, 17% concerned the training content, and 21% concerned the value to the trainee. Fifteen percent of apprentices did not comment on the item. Figure 11 presents a breakout of comments across the *VR trainers*, *trainees*, and *training content* categories.

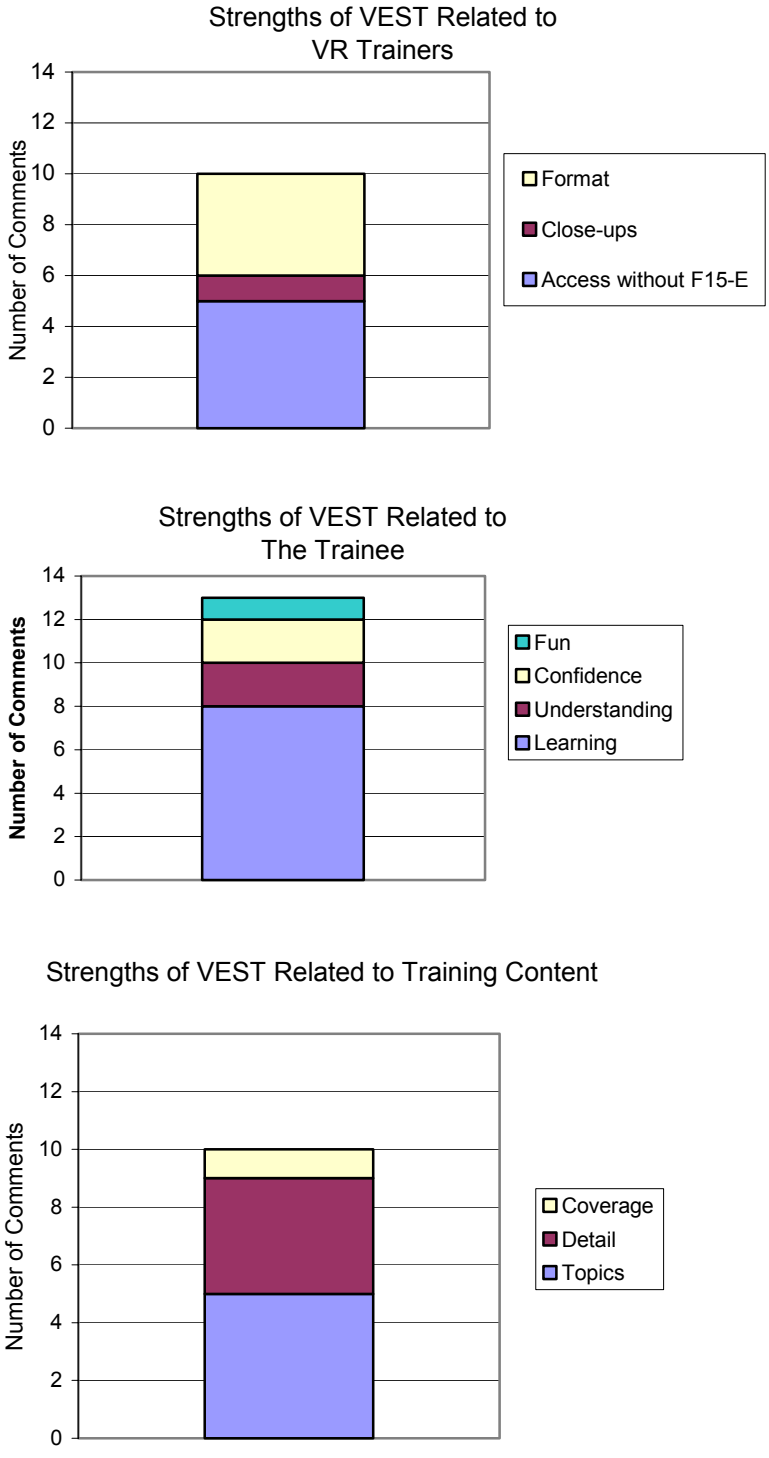

Figure 11. Strengths of VEST

Weak points of VEST. Ninety percent of apprentices commented on weak points of VEST. Of those comments, 6% of concerned the graphical user interface (blurry and slow in responding to movement), 1% concerned the lack of realism, 15% concerned the synthetic voice, 27% concerned adverse physical reactions, and 3% concerned the peripherals. Ten percent of apprentices did not comment on the item. Figure 12 presents a breakout of comments across the *VR trainers*, *trainees*, and *training content* categories.

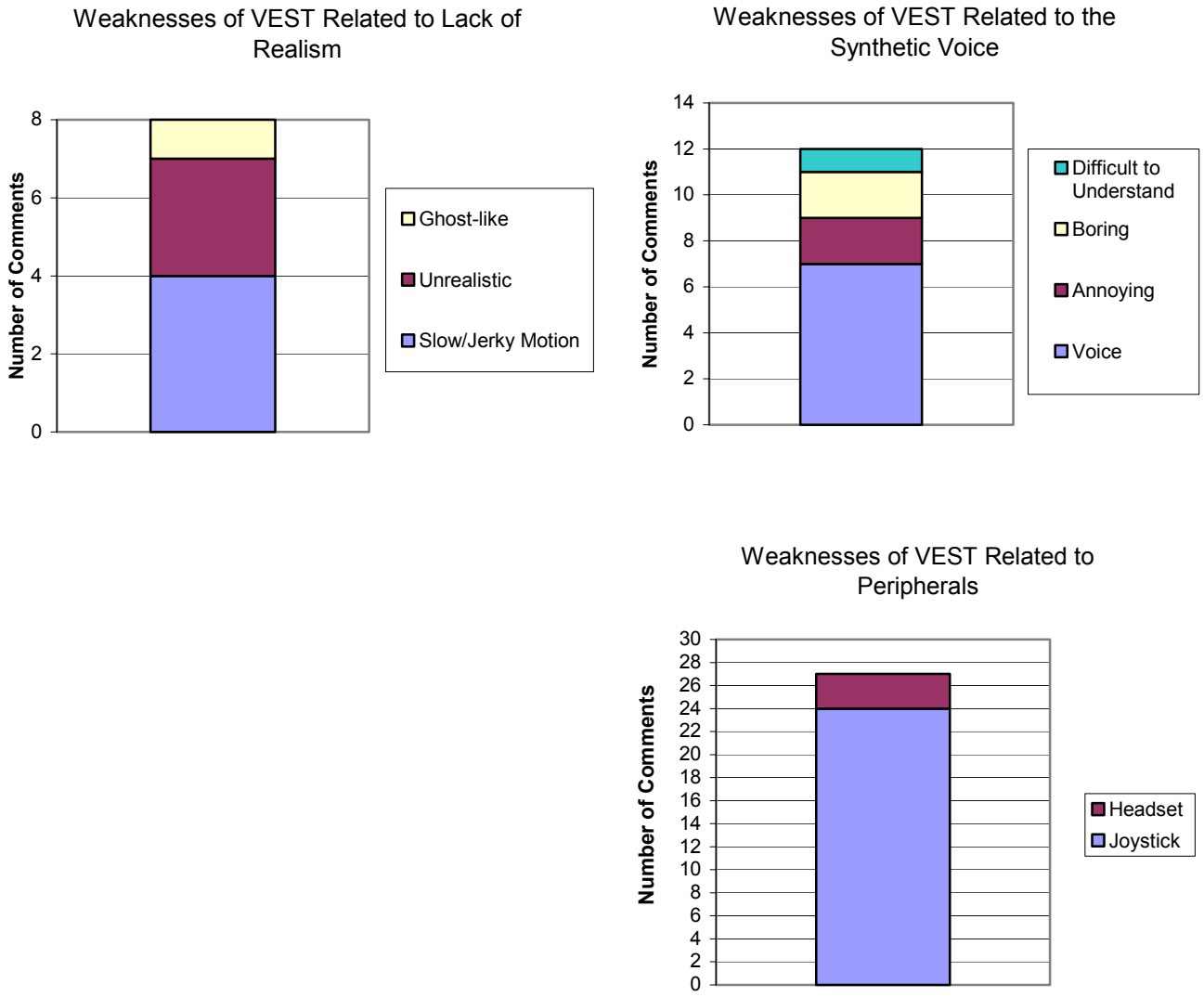

Figure 12. Weaknesses of VEST

Additional Comments. One-third of apprentices made additional comments. Fifty-four percent of apprentices who commented made positive comments and 31% made neutral comments. Two apprentices made negative comments. The negative comments were "*I think that VR is more of a game or entertainment environment not a training environment*." and "*The laser on the pointer should be adjustable. It made some operations almost impossible. I felt like Luke Skywalker trying to swing that thing in the cockpit*." All neutral comments addressed improvements to VEST. Figure 13 presents a breakout of comments within categories.

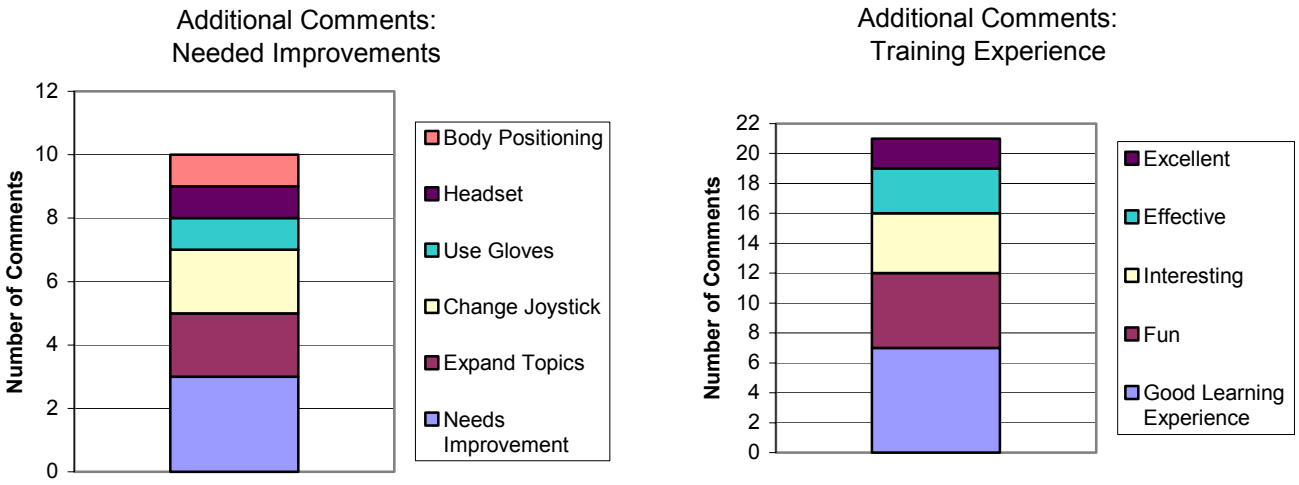

Figure 13. Additional Comments

Performance by Comment Category

A median split was conducted on overall performance scores to investigate differences in comment categories.

Patterns of comments on the "YES/NO" items were similar between the two groups created from the median split. Statistically, there were no group differences found across the survey items. Seventeen apprentices from the median and below group (n = 24) and 18 from the above median group (n = 25) reported no previous VR experience; 8 apprentices from the median and below group and 7 from the above median group reported they were unable to complete the VEST lessons without adverse physical side-effects; 3 apprentices from each of their respective groups reported insufficient realism with VEST; 5 apprentices from the median and below group and 3 from the above median group reported the voice was not understandable; 7 apprentices from the median and below group and 8 from the above median group reported difficulty with the peripherals (joystick, headset, system components); no one reported insufficient time to complete the class; and 13 apprentices from each of their respective groups reported VEST had no effect on the classroom training. Responses to recommending VR to others were the exception—7 apprentices from the median and below group and 4 from the above median group reported "NO" they would not recommend VR. Two apprentices from the median and below group did not respond to the item; this could be interpreted as not recommending VR.

The pattern of comment categories was examined between the two groups, as well. No statistically significant differences were found in comment categories between apprentices scoring at and below the median (89.6) and apprentices scoring above the median. However, group differences in the number of comments in the *strong points* categories approached significance ($\underline{X}^2_{(4)} = 8.36$, $\underline{p}$value = .08). Two noticeable differences in the number of comments in categories were 6 apprentices from the median and below group and only 1 participant from the above median group reported increased confidence/knowledge/fun as the strong points of VEST; and half as many apprentices from the median and below group (6) reported the instructional aspects of VEST as the strong points compared to the above median group (12). Example comments from the latter group are: "*Helps trainee learn more about the cockpit switches and their functions*," "*Familiarizes individual with the aircraft*," and "*Gives trainee the best understanding possible about the E model. Covers topics well*."

Discussion

The initial assessment of VEST provides considerable evidence for the value in training aircraft maintenance technician apprentices with VR technology. Although training in a VR environment may be a new experience for apprentices, their overall performance scores reveal acceptable levels of learning. Following is a synopsis of the apprentices' reactions to VR training:

- **VEST is well received overall as a training experience.**

  Nearly all apprentices (85%) report VEST having at least one strong point. The strengths fall into three general categories: the VR trainer (*it felt as if I was actually in the cockpit*), value to the trainee (*helps you understand the equipment and the job you are doing better*), and the training content (*lesson content was very detailed, great training*). Over half the apprentices react positively when given an opportunity to make additional comments. The effectiveness of the learning experience is often noted. Approximately one-third of apprentices continue to suggest improvements to VEST, as additional comments, signifying acceptance of the training technology for future use.

- **VEST positively affects traditional classroom training for some apprentices.**

  Nearly half of apprentices report a positive effect of VEST on their classroom training. The remaining half report VEST has no effect on their classroom training. Only two apprentices reported that VEST detracted from classroom training. The detraction was due in part to delays in accessing the system, as only two trainers are available at the 363d TRS.

- **Apprentices recommend expanding VR to other types of training.**

  Three-quarters of apprentices recommend VR training procedures that put personnel and equipment at risk and when there is not access to equipment on which they are to be trained, as exemplified with the F15 E-model at the 363d TRS.

- **Training with VR increases apprentice confidence interacting with the F-15E.**

  A majority of apprentices reported the VR environment was realistic enough that they would feel safe when entering the cockpit. However, apprentices' global comments about the realism of VR suggest they were not necessarily responding to the latter part of the question "feel safe entering the cockpit". Apprentices who comment positively to the item believe VEST provides sufficient cockpit familiarization and increases confidence around cockpit controls. Apprentices who comment negatively about system realism tend to prefer a hands-on learning style and articulate the need for improvements to visual quality and training content.

- **Lesson length is not satisfactory for many apprentices.**

  One-third of apprentices did not perceive the length of the VEST lessons as appropriate for learning. Perceptions of lesson length range from too long, too short, to some long and some short.

- **Various adverse physical effects are experienced in the VR environment.**

  Over half of the apprentices report experiencing at least one adverse physical reaction to VEST. Common adverse reactions are headache, nausea, dizziness, and eyestrain.

- **The peripherals (joystick and headset) and synthetic voice need improving.**

  One-third of apprentices report no difficulty with the peripherals and associated components (joystick and headset). However, the *qualifying* comments suggest that peripherals have usability issues. Changes are needed to the joystick, headset, and physical configuration relative to the seat. Additional comments suggest the synthetic voice needs improving, although, most apprentices (85%) report a *qualified* "YES" to the voice being easy to understand. They describe the voice as annoying, boring, and monotonic.

- **VEST needs improvements.**

  Over three-quarters of apprentices suggest changes to VEST. The areas of improvement cover hardware, realism, the synthetic voice, and graphical user interface. Nearly all apprentices (90%) report at least one weakness of VEST. These fall into four general *categories: lack of realism (need to improve the visuals), synthetic voice (the* voice was boring and monotone), adverse physical reaction (*I felt nauseous after the first couple of lessons*), and peripherals (*joystick too precise, headset uncomfortable*).

The assessment data allows researchers to begin considering the next generation specifications for an enhanced VR system by basing improvements to the system on apprentice performance and reactions and considering leading-edge VR technologies.

CONTINUOUS ASSESSMENT

Data Collection

The Virtual Environment Safe-for-maintenance Trainer is being continuously assessed. Apprentices are asked to enter comments about their experience at the bottom of their performance printouts, following completion of VEST. The analyses reported next cover performance data and comments collected from 84 apprentices assigned to complete VEST. Two apprentices were unable to complete VEST so their data were excluded from the analyses.

Performance Data

The table below contains the average performance scores across the VEST drills broken out by number of tries. The system failed to capture "aircraft safe for maintenance drill" scores for three apprentices, so their overall performance scores could not be computed. The overall average performance score was 80.3%. Only six apprentices had overall average scores below 70%.

Performance results indicated students had the greatest difficulty with the "left console aft cockpit switch drill." Even after the maximum number of tries (2) only seven apprentices were able to pass with acceptable scores (100%). "Review aft cockpit drill" and "weapons station drill" also indicated some degree of difficulty for 19 and 56 of the 82 apprentices, respectively. After two tries these apprentices scored only slightly over the 70% pass rate.

Practice had the largest impact on the "aircraft safe-for-maintenance drill" where those who required a second try at attaining a passing score outscored those who passed on a single try ($t_{(75)} = 3.34$; pvalue = .001) (Table 4).

Table 4. Average Performance Scores on VEST Drills by Number of Tries (n=number of apprentices)

	Number of Tries			Number of Tries	
Drill	1	2	Drill	1	2
Left console forward cockpit switch drill	81.6 (n=53)	81.1 (n=29)	Left console aft cockpit switch drill	---	25.6 (n=82)
Main console forward cockpit switch drill	82.0 (n=59)	82.1 (n=23)	Main console aft cockpit switch drill	91.4 (n=77)	90.0 (n=5)
Right console forward cockpit switch drill	90.0 (n=77)	82.6 (n=5)	Right console aft cockpit switch drill	84.4 (n=73)	81.2 (n=9)
Review aft cockpit drill	84.4 (n=63)	74.2 (n=19)			
Weapons station drill	99.9 (n=26)	77.7 (n=56)	Safety device drill	87.1 (n=73)	89.0 (n=9)
Aircraft safe for maintenance intro drill	91.6 (n=79)	95.5 (n=2)	Aircraft safe for maintenance drill	76.4 (n=22)	84.8 (n=55)

Apprentice Comments

The 84 apprentices made 263 total comments. One comment was extraneous ("Thanks VR!") so it was discarded. Comments were assigned a category and a valance (positive, negative, or neutral) by two raters.

Fourteen coding categories were used that encompassed all comments. Definitions of the categories follow:

Aircraft – comparison between VEST and F-15E
Content – reference to instructional material
Experience – reaction to VEST
Fix – suggestion to improve VEST

Headset – reference to VR headset

Learn – testified to learning something

Learning Style – expressed preference for an instructional method

Media – reference to training tool

Physiology – physical reaction to VEST

Pointer – reference to interface (hardware/software) for object selection

Unreliability – reference to system glitches

Usability – reference to interactions with VEST

Visual – reference to quality of visual experience/graphical representation of aircraft

Voice – reference to quality of aural experience

Cohen's index "kappa" was used to derive interrater reliabilities. Interrater reliabilities were .95 for category assignment and .90 for valance assignment. Both reliabilities are acceptably high. Raters came to an agreement on divergent category and valance assignments, after making initial ratings.

Apprentices on average made 3.1 comments (range 1-7). The average number of positive comments was 1.3 (range 0-4). The average number of negative comments was 1.4 (range 0-5). The average number of neutral comments was .4 (range 0-2). The comment distributions were all positively skewed, showing fewer people made increasing numbers of comments.

Table 5 presents the frequencies of comments by category and valance. Categories are divided in the table between *mostly* positive, *mostly* neutral, and *mostly* negative comments. Overall, 41.5% of comments were positive, 13.3% were neutral, and 45.2% were negative. The division of comment categories by valance reveals 107 positive comments distributed over 5 distinct categories and 106 negative comments distributed over 8 distinct categories. All comments (23) in the "Fix" category, except one, were assigned a neutral valance table).

Table 5. Number of Apprentice Comments by Category and Valance

MOSTLY POSITIVE			
Category	**Positive**	**Negative**	**Neutral**
Experience	**39**	**8**	**8**
Learn	**37**	**--**	**1**
Media	**18**	**2**	**1**
Aircraft	**8**	**--**	**--**
Content	**5**	**3**	**1**
COUNT	**107**	**13**	**11**

MOSTLY NEUTRAL			
Category	**Positive**	**Negative**	**Neutral**
Fix	**--**	**1**	**23**

MOSTLY NEGATIVE			
Category	**Positive**	**Negative**	**Neutral**
Headset	**--**	**5**	**--**
Learning Style	**--**	**12**	**1**
Physiology	**--**	**20**	**--**
Pointer	**--**	**27**	**--**
Unreliability	**--**	**9**	**--**
Usability	**2**	**21**	**--**
Visual	**--**	**10**	**--**
Voice	**--**	**3**	**--**
COUNT	**2**	**106**	**1**

No negative comments were made about learning or the value of having experience with an E-model aircraft. Example positive comments are "The VR program served its purpose very well and was extremely educating" and "gives you a realistic view of the aircraft and the loading stations". In turn, no positive comments were made about the peripheral devices, graphical representations, voice synthesis, physiology response, reliability, and learning styles. The negative comments are found in Table 6.

Table 6. Negative Comments within Comment Categories

Content
▪ It does not go into depth. ▪ My only complaint is it should have the entire cockpit, not big blank spaces where some panels are. ▪ Right wing pylon pizzo has some bugs to work out.
Experience
▪ It's helpful in a way but with the way it's set up some of the items are hard to get to, and sometimes if you can't get to an item you get frustrated and click on anything. ▪ If you get frustrated enough times you will fail. ▪ You have to follow the order the computer gives or you fail the exercise. That is what happened with me. ▪ It was a little too long. ▪ The cockpit did get a little boring. ▪ I did not like it. ▪ I didn't like VR!
Learning Strategy
▪ Didn't help as much as I thought it would because I am more hands-on. ▪ Overall, I think it would be better to have a cockpit trainer. I prefer hands-on. ▪ It's no exchange for hands-on work. ▪ Involves no hands-on experience ▪ I feel that the trainers would be better than VR. ▪ I feel that a cockpit trainer would have been a lot more useful. ▪ I think that cockpit trainer would have given more of a real world feel of the cockpit. ▪ The virtual maintenance can't compare with real maintenance. ▪ Not as good as the real thing. ▪ I would rather have a real E model here.
Media
▪ Instead of VR, it should just be computer-based. ▪ After completing this objective, I believe VR is a glorified video game.
Visual
▪ The VR didn't focus well. ▪ Had a little trouble seeing all the controls in the cockpit. ▪ It was hard to select buttons in the cockpit. No matter how much I adjusted the lens. ▪ Not clear enough visual-wise. ▪ The VR was not clear enough to read the buttons and switches. ▪ The clarity could have been a little better. ▪ It was hard to see the controls in the cockpit. ▪ Only problem, some buttons were hard to look at such as the trigger switch. ▪ On the Waldorf system you cannot focus in on the words unless you get really close to them.
Voice
▪ The voice was, at times, a little hard to understand. ▪ Get rid of the annoying computer voice. ▪ The animated voice is very monotone.
Unreliability
▪ As far as scoring these lessons it is not correct. ▪ The reliability of the whole system. It seemed whenever we were ready to use the system something was wrong. ▪ The program freezes up constantly. ▪ The VR "locked up" a lot and that was upsetting. ▪ Rizzo had trouble sometimes. ▪ The only problems I had with eh VR was the program itself, it kept locking up and stopping very often. ▪ One thing I didn't like was sometimes you click on the correct object and it says it's wrong. ▪ The computer kept malfunctioning and slow. ▪ The system did make me restart at unit 18 when I hit unit 21, and it made me do the SFM twice but did not record a final SFM score.

Peripherals

Headset

- The headset was not very well constructed.
- I tried to adjust focus but it didn't work.
- Headset did fit comfortably, but in the cockpit we should be able to move forward, and the smell of the headset made me sick.
- The headset was uncomfortable after using it or an extended period of time.
- The VR head unit is very uncomfortable.

Joystick (A.K.A. Pointer, Laser, Hand controls)

- Pointer was not very accurate.
- It's very hard to click on certain items (like the grounding cable pin hole).
- The laser was very precise and was hard to choose the right object.
- The only detail that I feel could be better is the accuracy of stick grip control and pointer.
- Sometimes the pointer is a little bit picky on where to select.
- The pointer was hard to use on some awkward spots, i.e., trigger.
- The pointer kept messing up which in turn brought my score down!
- The laser seemed to not pick up on some things causing me to seem like I missed it.
- I didn't like the pointer laser because it was really difficult to point at something without moving slightly and getting the wrong thing.
- The pointer was hard to aim where you wanted to click on and even if you were on the right switch it would not give you credit.
- The control stick caused a problem with getting questions correct/incorrect. I thought that the pointer could have worked much better. It was a pain having to keep messing with it to get it to work.
- My only problem in dealing wit the VR system was the joystick pointer. I found it very difficult to control and click on selected items.
- Some of the switches were hard to get to with the pointer. For example, the trigger in the rear cockpit was especially hard to select.
- The pointer was difficult to use. Trying to click on items that were small was frustrating. It was hard to move around with the laser pointer being so small. You can barely hold still long enough to see what you were clicking on and click on it too.
- Sometimes the pointer didn't register what you were clicking on.
- Make it easier to point to the desired object.
- The only thing that I didn't like was the pointer was very picky and hard to control.
- It was very hard to use the stick.
- The pointer was hard to work with.
- It was hard to use the pointer.
- It's hard to use the pointer on thin, small objects in VR.
- The only trouble I had was with the pointer. It was difficult to master.
- Hand controls, hard to click on items.
- It's really hard to click the trigger.
- Need to work on controls, too hard to click on exact item.
- The joystick is somewhat difficult to use.
- The trigger was not as effective as it could be.

Physiological Reaction

- The VR machine gave me some eye problems.
- It hurt my eyes.
- Caused my left eye to hurt as well as my head.
- Caused headaches at first before getting used to it.
- If not used to it can give you a bad headache.
- The vision creates headache.
- The graphics tended to give me headaches
- It would give me a headache after a little while, and I wear glasses and it was kind of awkward.
- It also gave me a headache.
- Too long on the VR gives you a headache.
- The only bad thing is it will give you a headache if you stay in too long.
- A couple of minutes through the lesson I started to get a headache and motion sickness.
- I did experience a little motion sickness.
- At times it made me feel sick.
- I know there isn't much that can be done about the light headed feeling or uneasy feeling, but....
- The last section made me dizzy.
- It gets you dizzy a little.
- It really disorients you.
- Sometimes it put me to sleep.

Usability

Interface

- My scores were affected by the glitches in the program.
- The navigation however was quite cumbersome.
- The hot spots were sometimes hard to click.
- It should also be easier to get to smaller switches and gun trigger.
- On the objectives where you have to make sure that some of the cockpit switches are in the right position, the program alternates between having to either click the switch to put it in the right position, or click the word off to get it in the right position. This confused me and caused me to have to retry the exercise.
- Even when you pick the correct item, sometimes it tells you you're wrong. So you go to something else and that's wrong because the first pick was correct but you didn't pick it in the exact spot it wanted you to.
- The cage and uncage seeker button, when you try to click on something that is right, the arrow barely moves, it will count you wrong.
- The controls were very inaccurate. After clicking the same object twice (which I knew was right) the program told me I was wrong, then highlighted the same object I had just selected.
- Some of the trig buttons and knobs are hard to click on and some you're not sure what they want you to do.
- Difficult to find pylons armament ground safety pin hole and able to click on it.
- Also, the download pin on the GUN.
- You have to be patient, because when you click on the button that they ask for and its wrong, the correct button highlighted is the same one you were clicking on.
- It was a little hard to click on some of the buttons in the cockpit.
- The computer tells you that something is wrong when you follow the job guide word for word.
- The SFM course was not very helpful when it came to reading the job guide.
- One thing I did not like was that the safety violations were often set off when I was not doing anything wrong.

Restricted Movement/Uncomfortable

- I didn't like the restricted movement in the cockpit.
- Wasn't too fond of the movement you were limited too.
- Nor did I like having to reach around the stick with my neck getting into an odd position.

Difficult to use

- I felt that the courseware was a bit difficult to use.
- It's difficult to work with.
- It was jerky.

Performance by Comment Category and Valance

A median split was conducted to create two groups of apprentices—those who performed at and below 50% of the class (average overall score of 80.5%) and those who performed above 50% of the class. Chi-Square tests revealed no difference in the types of comments apprentices made whether they performed at or above the median score or below the median score. Nor were differences found in the number of positive and negative comments across categories based on the median split.

No relationship was found between overall performance scores and number of comments made ($\underline{r}$ = .099). Apprentices were divided into three groups based on a comparison of the number of positive versus negative comments (+ > -; + = -; + < -). No significant differences were found in performance scores across the three groups [$\underline{M}_{+ > -}$ = 81.3 (n = 30); $\underline{M}_{+ = -}$ = 80.1 (n = 17); $\underline{M}_{+ < -}$ = 79.5 (n = 32)].

Discussion

Valuable information on the effectiveness and acceptance of VEST was communicated through the continuous assessment. The data show that VEST provides effective training and experiences that would otherwise not be afforded F-15E model aircraft maintenance apprentice technicians. Apprentice comments and performance scores reveal an overall positive learning experience with VEST. Their comments further reveal a negative perceptual experience. No relationships were found between performance and the number of comments made, comment category, or comment valance. This finding suggests that any negative reaction to VEST did not influence overall performance.

Improvements need to be made to VEST to minimize negative aspects of training in virtual environments. Learning in virtual environments does not have to be painful. The negative comments indicate “fixable” problems, e.g., graphics quality, joystick difficulties, monotonous synthetic voice, and general usability.

Some apprentices provide fixes for what they saw as problems within the learning environment. Example “fix” comments are: “Maybe if you could use a glove to touch an object and not use the joystick.” and “There is one thing that would help memorize switches and stations; that is to repeat the name of the station/switch after it is clicked in the introductory part of the lesson. “ and “I think if we also have a way to get the computer to further explain what it expects, that would help.” An in-depth examination of suggested fixes reveals 50% deal with replacing the pointer (joystick) with something less accurate and easy to use, like gloves. The remaining suggestions concern non-specific “needs improvement” (16%), instructional design improvements (16%), and visual quality improvements (16%), with one fix addressing light-headedness.

Negative comments about the VR experience target areas of improvement for both instructional and interface features of VEST. Additional research is required to better understand the physiological responses (e.g., headache, eyestrain, disorientation, nausea, muscle stress) of apprentices to the virtual environment to continually improve training effectiveness and acceptance. Future directions for VR training are discussed next.

GENERAL DISCUSSION

The Air Force Air Education and Training Command (AETC) has proposed the concept of a Generalized Operations Simulation Environment (GOSE) for aircraft maintenance training. The training system would meet the need for a common, cost-effective, generalizable VR training platform for aircraft maintenance. The new system would help address simulator shortfalls—non-concurrency, no availability, high life cycle and sustainment costs.

The Air Force Research Laboratory (AFRL) is taking on the challenge of creating GOSE for aircraft maintenance in collaboration with AETC. The way ahead for realizing GOSE involves reengineering VEST to a PC-based platform, using off-the-shelf software, developing instructionally sound authoring capabilities for VR environments, demonstrating the operational feasibility and military training utility of GOSE for aircraft maintenance. Results from the initial and continuous assessments help point the way from VEST to GOSE, beginning with necessary improvements to the instructional design and addressing usability issues raised from the assessment results.

Improvements to VEST

The **length of time a user is immersed** in the VR training environment is a concern. Although, apprentices were encouraged to take breaks every 15 to 30 minutes and not to stay in longer than 45 minutes, some chose to complete the training without breaks. **Adverse physical reactions** to the VR experience, such as headache and eyestrain, may be attributable to length of immersion. The instructional effectiveness of VEST may also be affected by the length immersion. There are reasonable points in training content for forced breaks that would shorten immersion time and support learning. They are cockpit switch familiarization, weapons station and safety device identification, and external aircraft safe-for-maintenance procedures.

Realism could be heightened by use of photographic images and different peripheral devices. Off-the-shelf VR applications have advanced to accepting embedded graphics (photographs) rather than constructed graphics in creating the virtual airframe. Reductions in motion lag time, making the airframe “solid”, and surround detail are needed enhancements to **sensory realism**. It is likely that such changes would diminish feelings of nausea and dizziness.

Peripheral devices, e.g., Pinch Gloves™ and action mat, are available to allow users to interact with the virtual airframe in manners similar to how technicians interact with an actual airframe. If the joystick is replaced with gloves, as suggested in the assessment, then users' experiences would more closely resemble hands-on training. Users would be required to reach out to interact with switches and ground maintenance safety pins. An action mat could be added as a peripheral when the user is required to walk around the airframe. Users could be seated when working in the cockpit. Regardless of the addition of the action mat, changes are required to eliminate interference of the current seat with the headset and joystick cables and **decrease user discomfort**.

The VEST system, like most other computer-based training systems, is not completely reliable. Performance data were lost for 12 apprentices in the initial assessment and 2 apprentices in the continuous assessment. A degree of unreliability is due in part to a lack of understanding of how to save data. There are training and technology solutions to saving data. The pointer, representing the joystick in the VR environment, was seen as unreliable. The pointer's extreme precision led to user frustration and failure at times. A switch from the joystick to Pinch Gloves™ could both **improve reliability** of the trainer and **increase realism**.

Instructional improvements are needed to (a) control for information overload, (b) expand content areas, (c) extend opportunities to practice, (d) guide learning, and (e) increase accessibility to the trainer. In addition, the ability to work collaboratively in the VR environment would enhance the training capability of the system.

Future Training Research in VR

Apprentices' reactions reveal many problems encountered while training in a VR environment. The problems fall into instructional and usability categories. Effectiveness of solutions to these problems is an empirical question. The Air Force Research Laboratory is working toward establishing a VR training research lab at Mesa to begin answering and asking instructional and usability questions about VR training environments.

The value of AETC's concept of GOSE needs to be determined for both training and operational units. GOSE's PC-based approach to immersive VR training makes the system affordable and accessible. The modular approach of GOSE is designed to make training available across airframes.

Follow-on assessment data is needed to determine the extent to which training with VR systems transfers to the operational environment.[1] Elements of the operational environment such as visual and auditory representations of the flightline in the VR environment would likely enhance transfer of training. There is much to discover through VR training research.

[1] See D. Jeffery, R. Greene, Levi, K., & Schneider, D. (2001). F-15E virtual reality interactive courseware simulation, armament maintenance training system. Proceedings from Interservice/Industry Training, Simulation & Education Conference. Orlando, FL.